City University of New York (CUNY)
CUNY Academic Works

2016

Free Energy-Based Virtual Screening and Optimization of RNase H Inhibitors of HIV-1 Reverse Transcriptase

Baofeng Zhang
CUNY Brooklyn College

Michael P. D'Erasmo
CUNY Graduate Center

Ryan P. Murelli
CUNY Brooklyn College

Emilio Gallicchio
CUNY Brooklyn College

How does access to this work benefit you? Let us know!

More information about this work at: https://academicworks.cuny.edu/bc_pubs/157
Discover additional works at: https://academicworks.cuny.edu

This work is made publicly available by the City University of New York (CUNY).
Contact: AcademicWorks@cuny.edu

This is an open access article published under an ACS AuthorChoice License, which permits copying and redistribution of the article or any adaptations for non-commercial purposes.

ACS OMEGA

Article

http://pubs.acs.org/journal/acsodf

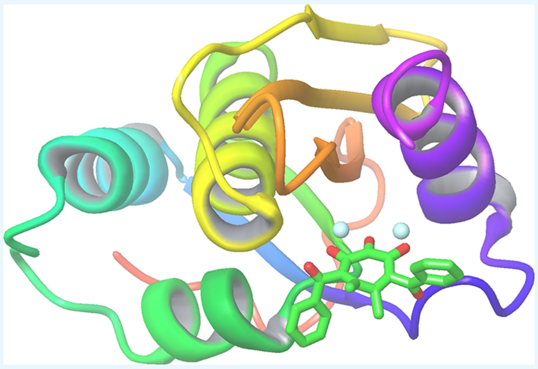
Free Energy-Based Virtual Screening and Optimization of RNase H Inhibitors of HIV-1 Reverse Transcriptase

Baofeng Zhang,[†] Michael P. D'Erasmo,[†,‡] Ryan P. Murelli,[†,‡] and Emilio Gallicchio*,[†,‡,§,∥]

[†]Department of Chemistry, Brooklyn College, City University of New York, Brooklyn, New York 11210, United States

[‡]Ph.D. Program in Chemistry and [§]Ph.D. Program in Biochemistry, The Graduate Center of the City University of New York, New York, New York 10016, United States

Supporting Information



ABSTRACT: We report the results of a binding free energy-based virtual screening campaign of a library of 77 α-hydroxytropolone derivatives against the challenging RNase H active site of the reverse transcriptase (RT) enzyme of human immunodeficiency virus-1. Multiple protonation states, rotamer states, and binding modalities of each compound were individually evaluated. The work involved more than 300 individual absolute alchemical binding free energy parallel molecular dynamics calculations and over 1 million CPU hours on national computing clusters and a local campus computational grid. The thermodynamic and structural measures obtained in this work rationalize a series of characteristics of this system useful for guiding future synthetic and biochemical efforts. The free energy model identified key ligand-dependent entropic and conformational reorganization processes difficult to capture using standard docking and scoring approaches. Binding free energy-based optimization of the lead compounds emerging from the virtual screen has yielded four compounds with very favorable binding properties, which will be the subject of further experimental investigations. This work is one of the few reported applications of advanced-binding free energy models to large-scale virtual screening and optimization projects. It further demonstrates that, with suitable algorithms and automation, advanced-binding free energy models can have a useful role in early-stage drug-discovery programs.

INTRODUCTION

It is very challenging to design potent and specific drugs for clinical use. The chemical synthesis of specific derivatives to probe binding preferences is often the most involved and time-consuming process. Information from experimental structures of receptor−inhibitor complexes, when available, is often an invaluable resource to guide the chemical synthesis efforts toward the most promising leads. Often, however, crystallographic data are limited to a very small fraction of chemical space and biological conditions. The design of human immunodeficiency virus (HIV)-1 RNase H inhibitors is a particularly difficult medicinal chemistry problem. The RNase H domain of the reverse transcriptase (RT) catalyzes the degradation of the DNA/RNA hybrid formed during the RT process.[1] Inhibition of this functionality of HIV RT prevents viral replication.[2] However, despite substantial efforts,[3−11] to this date there have been no clinically approved drugs that target the RNase H domain of RT. This is in contrast to the widely available nucleoside reverse transcriptase inhibitors[12,13] and integrase strand transfer inhibitors,[10,14,15] which target two-metal catalyzed nuclease functionalities similar to that of RNase H.

There is likely a fundamental biophysical basis for the lack of progress. The HIV RNase H active site is rather shallow and offers few specific structural anchors to exploit.[4,16] Compared to polymerization and integrase inhibitors, with half maximal inhibitory concentrations (IC_{50}'s) in the low nanomolar range, even the best RNase H inhibitors display relatively weak and nonspecific binding. The lack of specificity in turn causes toxicity due to unwanted binding to the structurally similar human RNase H and to other cellular enzymes. Lack of thorough structural and mechanistic understanding of the function of RNase H in the cellular context also poses additional challenges. For example, the effect of the RNA/DNA substrate on inhibitor binding is complex and poorly understood. Often RNase H inhibitors with promising in vitro characteristics do not display effective viral neutralization capacity when tested in vivo.[17,18]

Structure-based computer-aided drug design has become standard practice in drug-discovery programs in academia and industry. The fundamental idea is to use available crystallographic models to predict computationally the strength of binding of ligands to protein receptors to guide synthetic, biochemical, and medicinal efforts. Most often computational

Received: July 11, 2016
Accepted: September 7, 2016
Published: September 21, 2016

© 2016 American Chemical Society

DOI: 10.1021/acsomega.6b00123
ACS Omega 2016, 1, 435−447

modeling in this area is in the form of high-throughput virtual screens using fast docking and scoring methods capable of processing ligand libraries containing thousands of ligands.[19−23] Docking and scoring methods are particularly successful in screening out ligands unlikely to bind due to steric and energetic incompatibility and in providing structural models of the receptor−ligand complexes. They are, however, often unsuitable for accurate ranking of binders as well as for lead optimization. These applications are increasingly being addressed by physics-based methods that seek to directly compute the binding constant or, equivalently, the free energy of protein−ligand binding.[24] Relative free energy perturbation protocols aimed at estimating differences of binding free energies between related compounds have achieved a high level of reliability and automation.[25] Deployment of absolute binding free energy models within drug-discovery programs[26−32] is far less common. These are applicable to ligand libraries containing diverse scaffolds that are not amenable to relative binding free energy approaches. They are also suitable for comparative studies of binding to multiple receptors to, for instance, enhance specificity. From a biophysical point of view, the main advantage of binding free energy models is their ability to represent entropic and reorganization effects that oppose binding. We have shown for example that free energy models can improve screening enrichment of focused ligand libraries.[32,33]

In this work, we investigate computationally the binding of a library of α-hydroxytropolones (Figure 1),[34] a promising class of RNase H inhibitors,[17,18,35] by means of free energy modeling using the binding energy distribution analysis method (BEDAM) and advanced parallel molecular dynamics conformational sampling. The best synthetic α-hydroxytropolones characterized so far have exhibited in vitro potencies in the hundreds of nanometers range and a reasonable therapeutic window in cell-based antiviral assays. However, biochemical inhibition data collected so far do not show strong structure−activity relationships (SARs). The RNase H functionality of HIV-1 RT continues to be the subject of intense investigation because it is one of the few fundamental biochemical functions of the virus that is not yet targeted by antiviral drugs. Many inhibitors of HIV-1 RNase H have been discovered;[4,6,36] however, none of them has entered clinical trials due to poor in vivo activity. α-Hydroxytropolones, chiefly represented by the parent compound β-thujaplicinol, are observed to non-competitively bind to the active site of RNase H of HIV-1 RT. β-Thujaplicinol and manicol are two α-hydroxytropolone natural products, which have shown to potently inhibit HIV-1 RNase H with promising specificity and a favorable therapeutic window.[4,35] The α-hydroxytropolone core offers a convenient and flexible platform to build a large and diverse library of inhibitors. Our efforts are aimed at discovering synthetic α-hydroxytropolone derivatives with increased affinity and specificity for HIV-1 RNase H, so as to achieve enhanced in vivo potency. The modeling work reported here is part of an effort to gain further insight into the physicochemical properties of the RNase H active site to guide further synthetic and biochemical investigations toward this goal.[18]

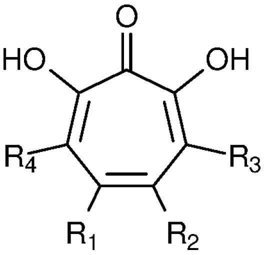


Figure 1. Nomenclature of side-chain substituents around the α-hydroxytropolone core.

This work also represents an interdisciplinary experiment aimed at evaluating the feasibility and usefulness of incorporating advanced-binding free energy modeling into the structure-based drug-discovery pipeline. Investigating SAR exclusively by experimental means, involving the synthesis of a large variety of derivatives and inferring their effects on receptor binding by measuring their in vitro and in vivo activities, is a tedious and expensive process. In addition, because of the complexities of inhibitory mechanisms, the relationship between receptor binding and activity is often unclear. The known structures of the RNase H domain of HIV-1 in complex with inhibitors, as well as with several α-hydroxytropolones, can serve as the basis of structure-based efforts to guide synthetic and biochemical efforts. However, standard docking and scoring computational approaches to predict affinities of α-hydroxytropolones have been of limited usefulness due to lack of discriminatory potential. Compounds known to have poor affinity are often docked and ranked as favorably as more potent ones. This is likely due to the shallow nature of the RNase H active site and subtle conformational reorganization effects, which affect binding.

Our absolute binding free energy model[29,37,38] is particularly suitable to tackle difficult systems such as this. It fully treats the thermodynamics of binding including the formation of ligand−receptor interactions and conformational reorganization processes. We describe in this work, several cases in which,

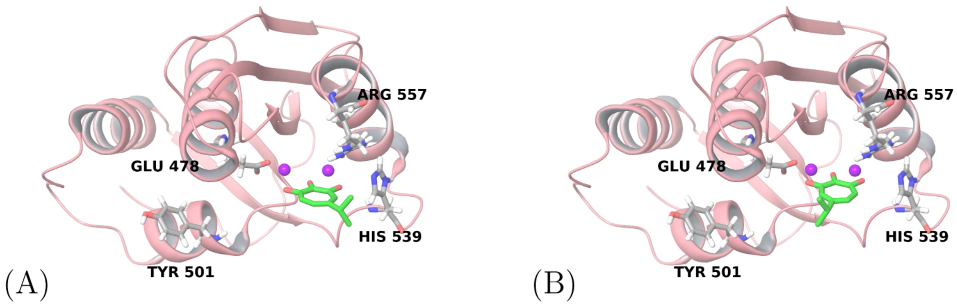


Figure 2. Crystal structure (3K2P) of the β-thujaplicinol bound active site of the RNase H (A), and β-thujaplicinol modeled in the corresponding flipped conformation, obtained by 180° rotation around the ligand axis in the plane of the bound cations (B).

for example, binding is enhanced by interactions with structural elements that form with some ligands but not with others. Conversely, in some cases, intramolecular hydrophobic interactions adversely affect binding. Representing these dynamical processes is key to understand binding propensities in this and other systems. At the same time, our methodology, because it treats each ligand independently and unlike double-decoupling methods,[39] requires only one thermodynamic transformation step and is straightforward to setup and deploy for relatively large ligand libraries. The protocol outlined here should find general applicability to structure-based drug-discovery programs.

■ MATERIALS AND METHODS

System Preparation. The computational model of the RNase H domain of HIV-1 RT was generated from the PDB crystal structure 3K2P[4] with bound β-thujaplicinol (Figure 2). Receptor preparation (hydrogen atom addition, protonation state selection, etc.) was carried out with the software "Protein Preparation Wizard, Schrodinger Suite (2014) LLC, New York, NY"[40] using the optimized potential for liquid simulations (OPLS) 2005 force field. Modeling of the ligand chelation of the magnesium cations[41] (replaced in the crystal structure by the corresponding manganese(II) cations) is important as these play a key role in the ribonuclease function of the enzyme.[4] The metal cations are known not to bind simultaneously to the binding site in the absence of a bound substrate or inhibitor,[4] and the alchemical process (see below) includes the modeling of the ligand-free state of the receptor. In addition, the applicability of classical nonpolarizable force fields, such as the one employed here, to describe interactions between organic ligands and divalent metal cations is challenging.[42,43] Hence, we opted to impose restraining potentials to the metal cations to maintain them bound to the active site residues. Specifically, flat-bottom harmonic potentials centered at 3.0 Å and of tolerance 2.0 Å were applied to the distance between metal ion A, which is the one closer to HIS 539, and the C_α atoms of residues 478, 443, and 498, and between metal "B" and the C_α atoms of the residues 443, 498, and 549 (metal cations are labeled as in ref 4).

We selected a virtual library of 77 α-hydroxytropolones (see Table S1), based on known leads and ongoing synthetic efforts.[17,18] β-Thujaplicinol, a naturally occurring α-hydroxytropolone inhibitor,[44] was set as the reference compound for the computation of relative binding free energies (see below). The classical force field used here is not suitable to describe accurately variations in the strength of the interaction between the ligand and the metal cations. The model employed here is meant to probe the effects of substitutions away from the metal chelation groups, assuming that metal chelation modality and strength do not vary significantly across the ligands in the library. Consequently, the other known natural α-hydroxytropolone inhibitor, manicol, could not be treated with the present model, given its unique mode of chelation of the metal cations.[17]

Ligand structures were prepared by using the LigPrep workflow within the Maestro program. The protonation and tautomerization states were generated at a target pH of 7.0 ± 2.0. The two hydroxyl groups on the tropolone ring, which contact the metal ions of the receptor, were modeled as deprotonated.[4,34,45] Ionization penalties were calculated at pH 7 with Epik[46] and added to the BEDAM results (see below) to yield protonation state-specific binding free energies. Stereoisomers resulting from alkyl nitrogen inversion, rotamers, and binding poses separated by high energy barriers were modeled individually, as they do not interconvert during the simulations. Motivated by the free energy combination expression,[31,47] the predicted binding free energy of a compound was taken as the most favorable among the ones obtained for all of its conformers and protonation states considered.

A total of 136 complexes resulting from multiple protonation states and alternative stereoisomers of the 77 ligands considered were input into the Glide docking program (Schrodinger, LLC). Glide receptor grid generation employed the structure of the receptor prepared as above. The crystallographic position of β-thujaplicinol in 3K2P was identified as the center of the docking search region. Positional constraints and metal coordination constraints were applied to correctly chelate the ligands with two Mn(II) ions. The resulting docked conformations of the complexes were used as initial structures for the binding free energy calculation.

BEDAM Binding Free Energy Protocol. The BEDAM[29,48] is an established protocol for the estimation of absolute binding free energies of molecular complexes.[31,33,49−53] The model employs the analytic generalized Born plus nonpolar (AGBNP2) implicit representation of the aqueous solution.[37] The standard binding free energy, ΔG°_{AB}, of a receptor A and a ligand B is expressed as[49]

$$\Delta G^\circ_{AB} = -k_B T \ln C^\circ V_{site} + \Delta G_{AB} \tag{1}$$

where C° is the standard concentration of ligand molecules (1 M, or, equivalently, 1/(1668 Å^3)) and V_{site} is the volume of the binding site. The first term of eq 1, $-k_B T \ln C^\circ V_{site}$, can be interpreted as the entropic work related to transfer the ligand from a solution at concentration C° to the binding site of the ligand−receptor complex, which depends only on the standard state and the definition of the complex macrostate and is independent of specific energetic and structural properties of the ligand and the receptor. The second term on the right-hand side of eq 1 is the excess contribution to the binding free energy. It represents the work to couple a receptor and a ligand, that is to turn on receptor−ligand interactions, while the ligand is confined within the binding site region. Formally, the excess binding free energy is defined as

$$\Delta G_{AB} = -k_B T \ln\langle e^{-\beta u(\mathbf{r})}\rangle_0 \tag{2}$$

where $\beta = 1/k_B T$, $\langle\rangle_0$ represents an ensemble average in the artificial state in which the ligand and the receptor are uncoupled (with the ligand confined within the binding site) when they interact only with the solvent continuum,[29] and $\mathbf{r}$ represents a specified conformation of the complex. In eq 2, $u(\mathbf{r})$ is the binding energy of the complex in conformation $\mathbf{r}$ defined as

$$u(\mathbf{r}) = U_1(\mathbf{r}) - U_0(\mathbf{r}) \tag{3}$$

where $U_1(\mathbf{r})$ and $U_0(\mathbf{r})$ are the effective potential energies of the complex when the ligand and the receptor are coupled and uncoupled, respectively.

ΔG_{AB} is evaluated by staged free energy perturbation[54] using the λ-dependent biasing hybrid potential energy function

$$U_\lambda(\mathbf{r}) = U_0(\mathbf{r}) + \lambda u(\mathbf{r}) \tag{4}$$

where λ is the free energy progress parameter, $0 \le \lambda \le 1$. ΔG_{AB} is defined as the free energy difference between the $\lambda = 0$ and 1 states of the complex. Other intermediate values of λ trace an alchemical thermodynamic path connecting these two states. Binding energy samples are collected from multiple molecular dynamic simulations at different values of λ. These are analyzed using the unbinned weighted histogram analysis method[52,55] to compute the binding free energy, ΔG_{AB}. A soft core binding energy function[31] is introduced to improve the convergence of the binding free energy near $\lambda = 0$

$$u'(r) = \begin{cases} u_{\max} \tanh\left[\dfrac{u(r)}{u_{\max}}\right]; & u(r) > 0 \\ u_{\max}; & u(r) \leq 0 \end{cases} \tag{5}$$

where $u_{\max}$ is a large positive value (1000 kcal/mol in this work). This modified binding energy function replaces the actual binding energy function (eq 3) wherever it appears, so as to cap the maximum unfavorable value of the binding energy while leaving the favorable value of binding energies unchanged.

The AGBNP2 implicit solvation model was used in this work. The full description of the AGBNP2 model is available elsewhere.[37] It is based on a geometrical-parameter-free analytical implementation of the pairwise descreening scheme of the generalized Born model.[56] The AGBNP2 model estimates the solvation free energy of the solute, ΔG_{solv}, in the form of

$$\Delta G_{\text{solv}} = \Delta G_{\text{elec}} + \Delta G_{\text{np}} + \Delta G_{\text{HB}} \tag{6}$$

where ΔG_{elec} is the electrostatic contribution, ΔG_{np} is the nonpolar contribution, and ΔG_{HB} is a solute−solvent hydrogen bonding term. The electrostatic term is calculated by means of a variation of the continuum dielectric Generalized Born model.[57,58] The nonpolar contribution includes cavitation- and dispersion-free energy terms; the first is modeled by a surface area model and the second by an expression based on the Born radius of each atom.[58]

To accelerate conformational sampling, a biased sampling and parallel Hamiltonian replica exchange method (HREM) was employed in this work.[29] In this scheme, pairs of simulation replicas periodically attempt to exchange λ values through a Monte Carlo λ-swapping move. Despite the use of HREM, we observed very slow interconversion between conformations of the ligands related by 180° rotation with respect to the metal coordination plane (Figure 2). To correct this deficiency in this work, we have individually considered the two orientations of the ligand. In this strategy, the binding free energy of each of the 77 ligands considered was estimated from the results of two BEDAM simulations, corresponding to the two orientations of the ligand, for a total of 272 BEDAM simulations. The binding free energy of the ligand is obtained by combining the BEDAM results for each conformation, assuming equal populations in solution of the two orientations.[59]

Orientations of β-thujaplicinol analogues are classified on the basis of the crystal structure of β-thujaplicinol in complex with HIV RNase H (PDB id 3K2P). In this structure, the isopropyl side chain is on the same side as the "A" metal cation and proximal to His 539 (Figure 2). For the analogues of β-thujaplicinol, we define the crystallographic orientation as the one in which the larger of the R1 and R2 side groups (see Figure 1) is at the same position as the side group of the isopropyl side group of β-thujaplicinol in the 3K2P crystal structure. The opposite orientation is classified as "flipped". Examples of crystallographic and flipped conformations for β-thujaplicinol are shown in Figure 2.

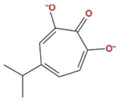
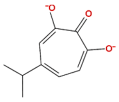
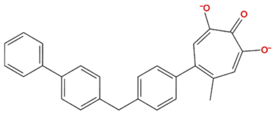
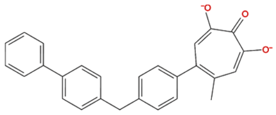
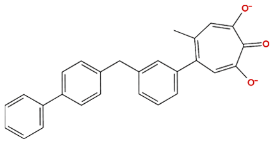
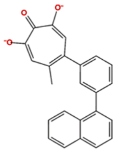
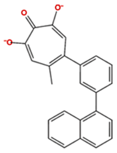
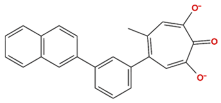
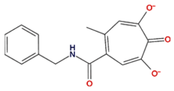
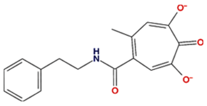
Computational Details. The BEDAM simulations employed the 2005 version of the OPLS-AA[60] force field together with the AGBNP2 implicit solvent model.[37] The force field parameters were automatically assigned by using Schrödinger's atomtyper.[61] The parallel alchemical Hamiltonian Replica Exchange molecular dynamics simulations were conducted using the Asynchronous Replica Exchange software[62] with the IMPACT MD engine.[61] Approximately half of the BEDAM free energy calculations were conducted on the SuperMIC cluster at the Louisiana State University as part of the Extreme Science and Engineering Discovery Environment (XSEDE) NSF consortium. The other half of calculations were conducted on the Brooklyn College WEB computational grid, based on the BOINC grid software. The simulation temperature was set to be 300 K for all simulations. A total of 18 intermediate alchemical steps at λ = 0.0, 0.002, 0.004, 0.008, 0.01, 0.02, 0.04, 0.07, 0.1, 0.17, 0.25, 0.35, 0.5, 0.6, 0.7, 0.8, 0.9, and 1.0 were used to conduct the λ-biased replica exchanges. The binding site volume was defined as the conformation in which the center of mass of the ligand was within 5 Å of the center of mass of the receptor, where the center of mass of the ligand was calculated including only the atoms of the core structure (C1-C2-C3-C4-C5-C6-C7). The center of mass of the receptor site was calculated in terms of the C_α atoms of residues 443, 444, 478, 498, 539, 549, and 557 (RT chain A numbering).[63] A flat-bottom torsional restraint potential in terms of the dihedral angle between the C1, C6 atoms of the ligand and the C_α atoms of the residue 478 and 443 was applied to confine the ligand in one of the two orientations relative to the metal coordination plane (see above). The simulation system was minimized and thermalized at 300 K. Bond lengths with hydrogen atoms were constrained using SHAKE. The mass of hydrogen atoms was set to be 5 amu. A 12 Å residue-based cutoff was imposed on both direct and generalized Born pair interactions. BEDAM calculations were performed for approximately 2.5 ns of MD simulation per replica (45 ns in total for each complex), starting from the docked ligand poses. Data from the last 2 ns per replica was used for free energy analysis. Binding energies were sampled with a frequency of 1 ps for a total of 18 × 2000 = 36 000 binding energy samples per complex.

■ RESULTS

Binding Free Energy-Based Screening of Ligand Library. The natural α-hydroxytropolone, β-thujaplicinol, was simulated starting from the 3K2P crystal structure.[4] The crystal structure of β-thujaplicinol in complex with the RNase H domain (3K2P) served also as a starting structural model for all of the complexes with α-hydroxytropolone derivatives considered here. β-thujaplicinol is known to coordinate two divalent metal ions at the RNase H active site through its carbonyl and hydroxyl functional groups.[4] In the crystal structure, the tropolone ring of β-thujaplicinol is approximately coplanar with the metal cations. The same coplanar conformation is seen in our molecular dynamics simulations of β-thujaplicinol and its derivatives in both the orientations considered (Figure 2). The pattern of interactions between the oxygen atoms of the tropolone ring and the metal cations is also reproduced. The isopropyl side group of β-thujaplicinol does not significantly interact with residues of the RNase H active site. Consequently, the binding free energy scores for the two binding orientations

DOI: 10.1021/acsomega.6b00123
ACS Omega 2016, 1, 435−447

Table 1. Computed BEDAM Relative Binding Free Energies, $\Delta\Delta G_b^o$, for Selected Inhibitors of HIV-1 RNase H

ligand id[a]	orientation[d]	ligand structure	$\Delta\Delta G_b^{ob}$
β-thujaplicinol	flipped		0 [c]
β-thujaplicinol	cryst.		0.8
23	cryst.		0.9
23	flipped		−2.0
24	flipped		−2.1
26	flipped		−2.3
26	cryst.		1.5
27	flipped		−4.6
42	flipped		−3.0
46	flipped		4.7

[a] See text for a description of ligand identifiers. [b] In kcal/mol. [c] Reference compound. [d] "Cryst." refers to the crystallographic orientation, and "flipped" refers to the flipped orientation.

are similar ($\Delta\Delta G_b^o$ = 0.8 kcal/mol, favoring the flipped orientation).

The BEDAM estimates of the relative binding free energies of the 77 analogues of β-thujaplicinol studied here are listed in Table S1 of the Supporting Information. The results for a selected set are shown in Table 1. Relative binding free energy estimates are expressed with respect to the binding free energy estimate of the flipped orientation of β-thujaplicinol.

Preference for Flipped Binding Orientation. As discussed above, to thoroughly score all possible binding modes, we separately evaluated the crystallographic and flipped binding orientations of each ligand (see Figures 2 and 3). Interestingly, for most ligands, the flipped orientation yielded a more favorable binding free energy. For example (see Table 1), ligand 23 in the crystallographic orientation has a relative binding free energy of 0.9 kcal/mol compared to −2.0 kcal/mol

DOI: 10.1021/acsomega.6b00123
ACS Omega 2016, 1, 435−447

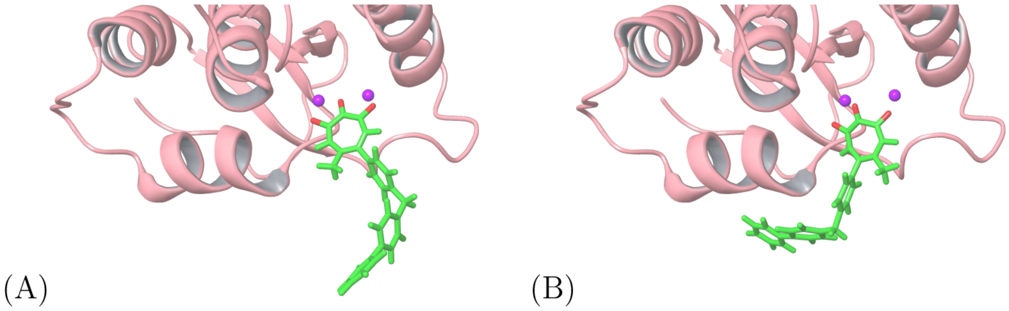
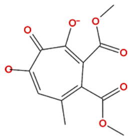
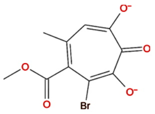
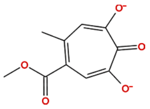
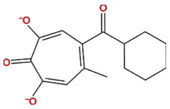
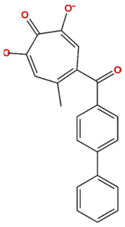


Figure 3. Representative structures of the crystallographic (A) and flipped (B) binding poses of ligand 23.

Table 2. Computed BEDAM Relative Binding Free Energies, $\Delta\Delta G_b^\circ$, and Experimental ΔT_m and IC_{50} for Some of the Inhibitors of RNase H

ligand id[a]	orientation[d]	ligand structure	$\Delta\Delta G_b^{ob}$	ΔT_m^c	IC_{50}^e
1	cryst.		−8.3	2.74	0.17
5	cryst.		−10.1	2.48	0.13
6	cryst.		−3.2	2.70	0.16
9	cryst.		−2.8	2.40	0.21
13	cryst.		−3.0	1.90	0.59

[a] See text for a description of ligand identifiers. [b] In kcal/mol. [c] In degree. [d] "Cryst." refers to the crystallographic orientation. [e] In μM.

for the flipped orientation, a difference of 2.9 kcal/mol favoring the flipped orientation. Similarly, the flipped orientation of ligand 26 is favored by 3.8 kcal/mol over the crystallographic orientation (Table 1). In general, we found that the flipped orientation is preferred over the crystallographic orientation in 61 of the 77 analogues.

Analysis of the molecular dynamics trajectories reveals that this behavior is due to the placement of the R2 side group within the interaction distance with the so-called "500" area[64] of the RNase H active site. Illustrative structures for ligand 23 are shown in Figure 3. In the crystallographic orientation (Figure 3A), the long aromatic R2 substituent of the ligand does not strongly interact with neighboring residues. Conversely, in the flipped orientation (Figure 3B), the R2 substituent forms strong hydrophobic interactions with GLN 500 and TYR 501. These interactions stabilize the complex and lead to a more favorable binding free energy.

The ratio of populations of the two binding poses is equal to the ratio of the Boltzmann factors of their corresponding binding free energies.[31,59] For example, the calculations predict that the flipped binding pose of ligand 23 is approximately 190 times more probable than the crystallographic pose. In contrast, β-thujaplicinol, for which the difference in the binding free energy between the two orientations is significantly smaller, is predicted to position in the two orientations with similar probability. This is consistent with the crystallographic structures.

Biochemical evidence and SAR analysis suggest that in the cellular environment, in the presence of the RNA/DNA hybrid substrate, the flipped binding pose may not be accessible to α-hydroxytropolones inhibitors.[65] It is believed that the substrate, anchored to the polymerization site[65] and the inhibitor, bound to the RNase H active site, can bind concurrently to the HIV RT dimer. Furthermore, according to this model, the 3′ end of the RNA strand covers the 500 region of the RNase H domain[65] and sterically limits the orientation of the inhibitor. Whenever useful, below we analyze the modeling results for each orientation of the ligand individually. The significance of the crystallographic orientation rests on its potential relevance for in vivo inhibition. The flipped orientation is significant for its potential to acquire additional inhibition activity by means of substrate displacement and as a starting point for the design of more potent and specific inhibitors incorporating DNA intercalators.[66]

Effect of R3 Side Group. The most evident SAR emerging from the binding free energy data is the significantly higher affinity of compounds substituted at the R3 position. For the reference compound, β-thujaplicinol, and for most of the ligands in the library, R3 is hydrogen. In compounds 1 through 5, this position is replaced by methyl and phenyl esters, methyl, or bromide (Tables 2 and S1). All of these substituents at R3, although less so for methyl, result in significantly more favorable binding free energies than β-thujaplicinol (gains as high as ~10 kcal/mol) in both the crystallographic and flipped poses. The improvements in binding free energy is likely overestimated. Nevertheless, the qualitative trend appears to be robust based on the structural information gained in the simulations. In the crystallographic orientation, the substituent at R3 fits into a side pocket lined by ARG 557, ALA 538, and HIS 539. Interestingly, we found that ARG 557, which is normally disordered, in the presence of a substituent at R3 forms a strong solvent-shielded hydrogen bond with one of the deprotonated hydroxyl groups of the ligand. In addition, see, for example, Figure 6, this interaction is further strengthened by hydrophobic interactions between the arginine head group and the aliphatic end group of the R3 substituent. In the flipped conformation (see, for example, Figure 7), the R3 substituent occupies a pocket on the other side of the active site, forming very stable interactions with TYR 501 and SER 499. Similar affinity gains are predicted for R3 substitutions applied to the crystallographic or flipped poses.

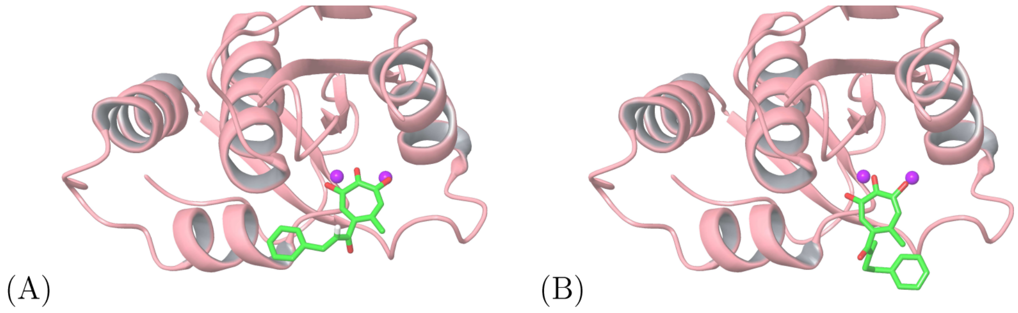
Effect of R2 Side Group. The greatest chemical variation in the ligand library investigated is focused on the size and nature of the R2 side group. The reference compound, β-thujaplicinol, has an isopropyl R2 side group. All of the members of the library, except β-thujaplicinol, have a methyl substituent at R1. We have investigated mostly aromatic substituents at R2, linked to the tropolone ring by means of alkyl, aromatic, piperidine, carbonyl, and amide linkers (Table S1). In contrast to the R3 side group, we have generally observed a weak dependence of binding affinity on the composition of the R2 group.

The effect of R2 is especially weak and often unfavorable for the crystallographic binding orientation. About two-thirds of the compounds unsubstituted at R3 have a worse predicted binding affinity relative to β-thujaplicinol when in the crystallographic orientation, and only three compounds (ligand 6, 9, and 13) have a predicted binding free energy gain of 2 kcal/mol or better (the assumed accuracy limit of the computational model). The best gain is −3.2 kcal/mol achieved for compound 6 with a methyl ester substituent at R2 (Table 2). This a relatively small gain compared to those for substitutions at R3, which are as high as 10 kcal/mol. The best compounds in this category, 6, 9, and 13, have an aromatic or alkyl side group at R2 joined to the tropolone ring by a carbonyl linker. We found that in the best binding poses, the R2 head group makes hydrophobic interactions with ALA 538 and the C_β of HIS 539. Compound 13, which has a longer R2 side group, makes additional interactions with TRP 535, PRO 537, and less often with LYS 540. These interactions are possible only when selecting an appropriate rotameric state for R2. We observed that rotation of the carbonyl linker is highly restrained (especially when R3 is substituted) and is rarely observed in the simulations. In a number of these circumstances, we have scored various R2 and R3 rotameric states individually (see below).

Apart from the moderate gains noted above, substitutions at R2 are predicted to have small or negative effects on affinity when in the crystallographic orientation. The R2 position points away from the body of the receptor, so opportunities for strong ligand−receptor interactions are few. In many cases, the R2 substituent is found to be mostly solvated and not contributing significantly to binding. Polar substituents tend to weakly oppose binding due to partial desolvation. This negative trend is amplified for charged substituents (for ammonium-containing groups in particular), for which binding is predicted to be significantly weak. Larger hydrophobic groups, except for the case noted, tend to disfavor binding, likely because of unfavorable desolvation of HIS 539, which is often seen to be turned away from the binding site. Interestingly, particularly long R2 groups, such as for ligand 64 (Table S1), are seen to loop around to make contact with TYR 500 and nearby residues, which are normally within the interaction distance only when the tropolone ring is in the flipped pose. However, these interactions are weakened by intramolecular strain of the R2 side group, as evidenced by the lack of similar conformations in the unbound conformational ensemble of the ligand. In other cases, large and flexible hydrophobic groups form strong intramolecular interactions, which further disfavor binding (discussed below).

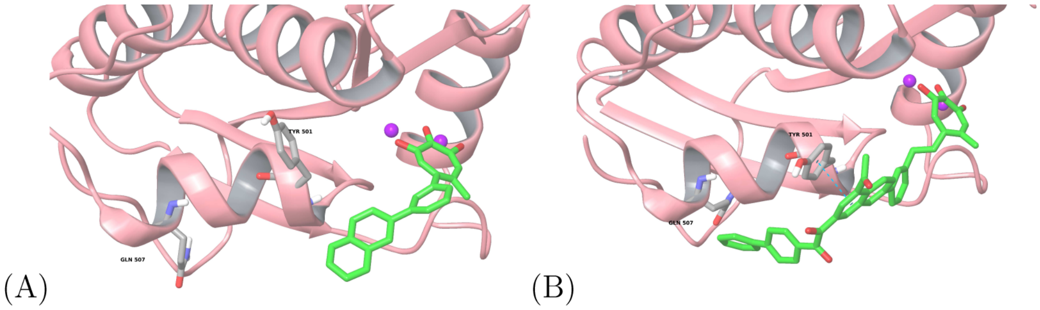
The R2 side group has a greater effect on binding affinity in the flipped binding pose. R2 side groups containing piperidine linkers generally lead to poorer activity, especially when the nitrogen is protonated. This effect does not seem to be directly related to the piperidine group itself but rather to the hydrophobic head and greater length of the R2 side group that contain it. The piperidine group tends to interact with TYR 501, whereas the hydrophobic head group is forced to extend into the solvent. With other linkers, nitrogen-containing aromatic head groups tend to lead to better affinity but only if not protonated. Ligand 27 (see Table 1), where R2 is a phenyl linker and a naphthyl head group, is the compound predicted to have the best affinity among those lacking a R3 substituent. The R2 side group, in the flipped pose, interacts with GLN 500, TRP 535, and PRO 537. To do so, the R2 group is required to assume a specific rotameric state. Ligand 27 has been chosen as the starting point for focused optimization below. As exemplified by ligand 20 (see Table S1), in which R2 is a biphenyl group, interactions of R2 with the 500 region of the receptor can be established even with relatively rigid moieties. In this case, it is the tropolone ring that tilts away from the optimal metal coordination plane to allow optimal positioning of the R2 group near TYR 501 and GLN 500.

In general, among the ones we considered, the nature of the R2 side group linker is predicted to be a minor determinant factor of improved affinity in both the crystallographic and flipped binding orientations. Ligands 23 and 24, for example, which differ by the para- or meta-substituted phenyl linker (see Table 1), have very similar predicted binding free energies (−2.0 and −2.1 kcal/mol, respectively) in their flipped binding orientations. Ligands with similar head groups having either



(A) (B)

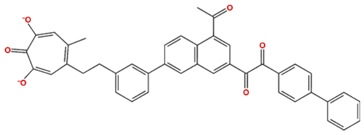
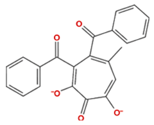
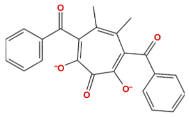
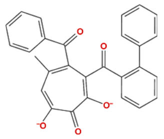
Figure 4. Representative structures of the complexes with ligands 42 (A) and 46 (B). The binding of ligand 46 is disfavored by intramolecular hydrophobic interactions.



(A) (B)

Figure 5. Representative structures of the complexes with ligand 27 (A) and R2-optimized ligand 78 (B).

Table 3. Computed BEDAM Relative Binding Free Energies, $\Delta\Delta G_{\mathrm{b}}^{\circ}$, for Optimized α-Hydroxytropolone Inhibitors of RNase H

ligand id[a]	orientation[c]	ligand structure	$\Delta\Delta G_{\mathrm{b}}^{\circ}$[b]
78	cryst.		−6.6
79	cryst.		−10.7
80	cryst.		−16.6
81	cryst.		−10.4

[a] See text for a description of ligand identifiers. [b] In kcal/mol. [c] "cryst." refers to the crystallographic orientation.

amide or carbonyl linkers are predicted to bind with similar strength.

Intramolecular Hydrophobic Interactions Decrease Binding Affinity. The computational model employed in this work attempts to take into account the conformational variability of both the ligand and the receptor. This feature has proven to be useful in the rationalization of surprising trends in binding affinities. For example, even though they differ by only one additional methylene linker group, we predict a large difference between the binding affinities of ligands 42 and 46 (see Table 1). In this case, and in a few other cases, we observed that, during the molecular dynamics simulations, intramolecular interactions develop between side groups of the ligand, which adversely affect the binding affinity. As illustrated in Figure 4A, and consistent with its favorable binding free energy, ligand 42 forms favorable interactions with the 500 region of the receptor through its R2 substituent. Conversely, due to its greater length and flexibility, the R2 substituent of ligand 46 is found to loop back to form a strong intramolecular hydrophobic interaction with the methyl group at the R1 position of the tropolone ring. Hence, the R2 side group of ligand 46 is not available to interact favorably with the receptor. In addition, the bulkier ligand volume shields the highly polar active site from the solvent, thereby decreasing the effect of favorable dielectric polarization. Both of these effects contribute to the poor predicted binding affinity of ligand 46 relative to ligand 42 despite their chemical similarity. Examples such as this underscore the complexities of molecular interactions and the difficulties in identifying SARs of general applicability.

Binding Free Energy-Based Virtual Lead Optimization. Some of the best compounds emerging from the free energy-based screening of the library were chosen as the starting point for the optimization attempts of side groups. We focused on the optimization of the R2 side group (see Figure 1 for nomenclature of side groups), using ligand 27 as a starting point (Table 1), and on the R3 side group, using ligand 1 as a starting point (Table 2).

Optimization of R2 Side Group. Ligand 27 with the benzyl-naphthyl substituent at the R2 position showed the best affinity among the 72 members of the library lacking a R3 substituent (see Figure 5A and Table S1). As for most other ligands in this class, ligand 27 shows a preference to bind in the flipped conformation, where the side group forms interactions with residues proximal to TYR 501. The optimization strategy we followed consisted therefore in extending the head groups to the R2 side chain of ligand 27 to form interactions between the R2 side groups and the residues along the 500 region of the receptor, hypothesized to play a role in the allosteric inhibition of RNase H.[67] The R2 side group was modified in a sequence of steps, which included extending it with biphenyl and ethyl linkers, and derivatizing it with acetyl and dicarbonyl polar groups to improve solubility and to attempt to form hydrogen bonds with the TYR 501 and GLN 500 residues. The optimization results are interesting. The final optimized inhibitor (ligand 78, see Table 3) does provide better binding affinity, having a binding free energy of −6.6 versus −4.6 kcal/mol for ligand 27. From Figure 5B, it can be seen that a $\pi-\pi$ stacking interaction is formed between the head group of TYR 501 and the naphthyl ring; however, no strong hydrogen bonding is observed between the carbonyl linkers and the residues along the 500 helix. As expected, the added biphenyl head group accesses the area of GLN 507 residue, which was predicted to be a promising binding region for inhibitors.[64]

Optimization of R3 Side Group. As mentioned above, the molecular dynamics free energy calculations revealed that substitutions at the R3 position (ligands 1 through 5) could induce recruitment of ARG 557 to form strong ionic interactions with one of the deprotonated hydroxyl of the α-hydroxytropolone ligand. In the absence of the R3 substituent, ARG 557 is consistently disordered and does not contribute to binding. This observation motivated us to consider the R3 position as a useful optimization target.

Of the five ligands in the library with substitutions at R3, we chose ligand 1 with a methyl ester substituent as a starting point because of the opportunity of optimization of hydrophobic interactions with the side chain of ARG 557. Replacing the methyl ester group with a phenyl-ketone group, which is thin enough to fit into the groove between ARG 557 and HIS 539 and at the same time has the potential to form $\pi-\pi$ interactions with the head group of ARG 557, increased affinity by approximately 2.4 kcal/mol. We observed that the resulting compound (ligand 79 in Table 3) was also able to form transient interactions with HIS 539. In an attempt to stabilize the latter interactions, we introduced a second phenyl group in the ortho position relative to the carbonyl linker. This second alteration (ligand 81 in Table 3) indeed resulted in the formation of stable $\pi-\pi$ interactions with HIS 539 and increased the affinity by 2.1 kcal/mol (Figure 8).

It should be noted that ligand 1 and its descendants (ligands 79 and 81) display severely hindered rotation of the ester side groups at the R2 and R3 positions. Indeed, interconversion between rotamers may be slow (on the order of hours or more) so as to be considered separate atropisomer species.[68] Rotameric states of these ligands have been simulated individually, as they do not interconvert during the molecular dynamics trajectories. In particular, ligand 79 has four atropisomer-like states, all of which have been simulated and only one (shown in Figure 6) displaying strong affinity to RNase H (−10.7 kcal/mol). Only in this particular conformation, in which the phenyl groups of the R2 and R3 side groups point toward opposite sides of the tropolone ring, can the ligand effectively engage ARG 557, as described.

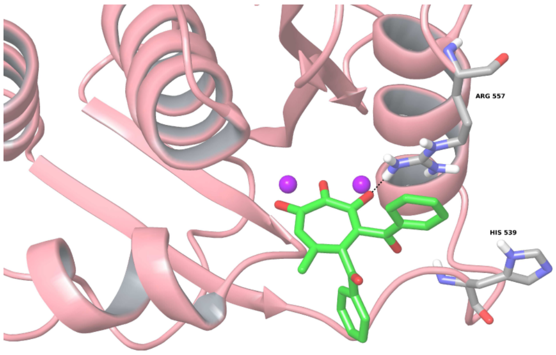


Figure 6. Representative structure of the complex with ligand 79. The HIS 539 and ARG 557 residues are shown in licorice representation.

Addition of R4 Side Group. As noted above, most ligands in the library, including those with substitutions at R3, bind more strongly in the flipped orientation. In the flipped orientation, the R3 side group is accommodated in a side pocket lined by GLN 500 and SER 499. Due to steric constraints, this binding mode is accessible only by some of the atropisomers of the ligand (see, for example, Figure 7). To probe the benefits of forming ligand−receptor interactions on both sides of the tropolone ring, we have tested the possibility of adding substituents both at R3 and at the symmetric opposite position, which we named R4, resulting in ligand 80 (Figure 7). In this ligand, the R3 and R4 symmetric positions are occupied by

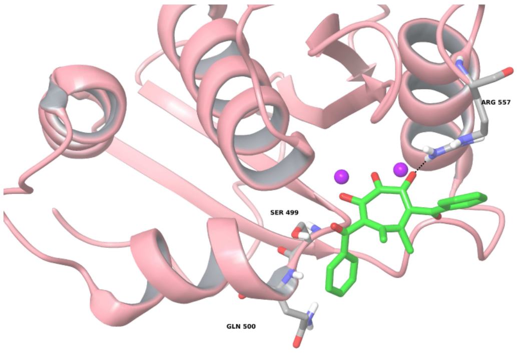


Figure 7. Representative structure of the complex with ligand 80. The SER 499, GLN 500, and ARG 557 residues are shown in licorice representation.

phenyl-ketone groups and the R1 and R2 symmetric positions by methyl groups. The two methyl groups are responsible for the atropisomerism, as they hinder free rotation around the carbonyl-tropolone ring. The specific atropisomer capable of binding has the phenyl groups on opposite sides of the tropolone ring (Figure 7). The predicted gain of affinity over β-thujaplicinol is very significant (−16.6 kcal/mol), albeit most likely overestimated. One of the phenyl-ketone groups is found in the small pocket lined by GLN 500 and SER 499. Although this group occupies the general region occupied by the R2 group in the flipped pose, it is far more sterically constrained. Hence, we predict that strong inhibitors will have asymmetric R3/R4 substitutions in which one group will be fairly small (a phenyl-ketone substituent being close to the maximum size), whereas the other will be able to assume a variety of sizes and shapes.

■ DISCUSSION

Modeling work in the context of medicinal chemistry applications,[69−71] including RNase H inhibition,[3,72] is most often based on molecular docking and scoring. We have recently shown that large-scale free energy-based ligand screening, while computationally more involved, can offer greater discriminatory potential in challenging cases like the present.[33] Due to the rather feature-less binding site, in this work, docking scores did not show clear binding trends. Starting from structures predicted by molecular docking, we have scored the binding of the ligands in the library by a binding free energy computational protocol, which yields relative thermodynamic dissociation constants directly comparable, in principle, to experimental measurements.

The greater resolution power of binding free energy models rests on their capacity to model reorganization and entropic thermodynamic driving forces not easily accessible by other means. Reorganization refers to the energetic and thermodynamic changes that accompany the remodeling of the conformations of the ligand and the receptor to enable binding. This process, often referred to as "induced fit" or "conformational selection", is characterized by free energy changes caused by intramolecular strain and entropic losses, that oppose binding. In addition, as here, the enzyme−ligand complex can have substantial conformational variability and cannot be fully described by any single representative conformation. Conventional models of binding based on static structures, which do not incorporate these thermodynamic signatures, are less likely to fully discriminate favorable from less favorable chemical modifications[33,53] and are therefore less informative in difficult cases like the present.

The binding free energy calculations have identified the R3 side group as the most promising handle to achieve high-affinity binding. The receptor pockets predicted to be occupied by the R3 groups are not apparent in the crystal structure of RNase H bound to β-thujaplicinol. When considering the crystallographic orientation, binding of the R3 group is predicted to critically depend on the recruitment of ARG 557, which is otherwise removed from the binding site region. Similarly, placement of the R3 group in the SER 499-adjacent pocket follows an overall expansion of this region and repositioning of side chains to optimize ligand−receptor interactions. The thermodynamic consequences of these and other key reorganization processes are taken into account by the model to discriminate favorable modifications, to be targeted by further investigations, from less promising ones.

The ligand design opportunities at the R3 position indicated by the present computational modeling efforts are supported by the available experimental data. Murelli et al. have measured the potencies of ligands 1 through 22 in terms of receptor thermal shift (ΔT_m) and IC_{50} values.[18] In that study, it was observed that, whereas increasing the size of the R1 substituent had consistently negative effects on activity, substitutions at R3 could accommodate a range of side groups without loss of affinity. For example ligands 5 and 6 (Table 2), which differ only in the presence of the bromine atom at R3, have similar ΔT_m and IC_{50} values (2.48 vs 2.70 °C and 0.13 vs 0.16 μM, respectively). Ligand 1, which has a bulkier R3 substituent, is as potent as ligand 5. Hence, both calculations and experiments indicate that the R3 position can represent a useful handle for the design of side groups leading to increased potency. The simulations identified that the R3 side group induces the formation of a binding groove sandwiched between ARG 557 and HIS 539, the latter being a key participant to the nuclease function of RNase H. Specific modifications of R3 are predicted to lock HIS 539 through strong electrostatic and π−π interactions, as seen, for example, in the optimized ligand 81 (Figure 8). The simulations have also indicated substantial potential benefits in adding substituents at the symmetric opposite position to R3 (the position labeled R4 in Figure 1) as seen, for example, in the optimized ligand 80 (Figure 7). Doing so accesses a small pocket near the TYR 501 and SER 499 residues. It is unclear whether the pocket is actually available in the presence of a substrate (see below); nevertheless, it might be worthwhile to investigate small substituents such as methyl or bromide at this position.

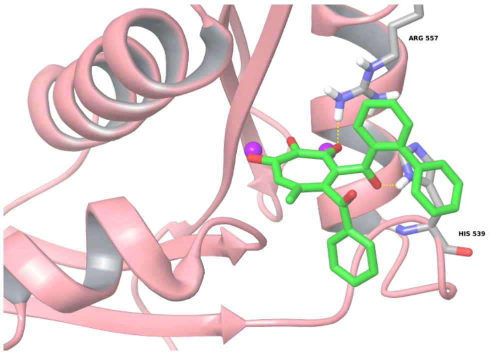


Figure 8. Representative structure of the complex with ligand 81. The HIS 539 and ARG 557 residues are shown in licorice representation.

Crystallographic structures[4,7] show that α-hydroxytropolones, similar to other RNase H active site inhibitors,[8−10] coordinate the two metal ions with the tropolone ring placed in the same plane as the metal ions. This configuration is preserved in the simulations of all of the ligands we investigated. The metal coordination geometries and the approximate C2 symmetry of the α-hydroxytropolones allows for two distinct binding modes, related by a 180° rotation (Figure 2), which were individually investigated for each ligand. In general, stronger affinity was obtained when the ligand was flipped relative to the crystal structure of β-thujaplicinol and manicol. When considering the flipped binding orientation, many ligands in the library are predicted to bind better than the reference compound (β-thujaplicinol). This is in contrast to the recently acquired experimental inhibition data on 22 of the 77 ligands investigated here,[18] which showed relatively small variations in potency which were for the most part unfavorable relative to β-thujaplicinol. The observed binding trends are reproduced when flipped binding modes are excluded from the computational model. When both binding orientations are considered, a majority of ligands are predicted to bind significantly better than β-thujaplicinol (17 out of 22 with a −2 kcal/mol cutoff, see Table S1). When, however, only the crystallographic orientation is considered, the range of binding free energies is substantially reduced, and only eight ligands show significantly better binding than β-thujaplicinol. Overall, these comparative data suggest that, under the conditions of the experimental inhibition assay, the flipped conformation is not available to the inhibitors.

The computational model does not take into account the nucleic acid substrate present in the inhibition assay. It is known[35] that β-thujaplicinol displays noncompetitive inhibitory behavior consistent with the hypothesis that an inhibitor and a substrate can bind the enzyme at the same time, possibly forming direct interactions, as for example in the homology model shown in Figure 9. This structural model is consistent with the observation that modifications at the R1 position are severely limited sterically and that modifications at R2 (more removed from the assumed position of the substrate) are also sterically disfavored but to a smaller extent.[18] According to this structural model, the flipped binding mode is not accessible by the ligands, as it would cause clashes with the substrate. Hence, the computational data obtained here, which are consistent with experimental evidence only when the flipped binding mode is ruled out, further support the model that displacement of the nucleic acid substrate from the active site does not cause dissociation of the substrate, which remains in the proximity of the RNase H active site. The model also explains why substitutions at the R3 position, which is the farthest from the substrate, can accommodate a larger variety of side groups. The evidence collected here suggests that inhibitor design strategies would benefit from taking into account interactions with the substrate, and that inclusion of the nucleic acid substrate in computational models would lead to more reliable computational predictions.

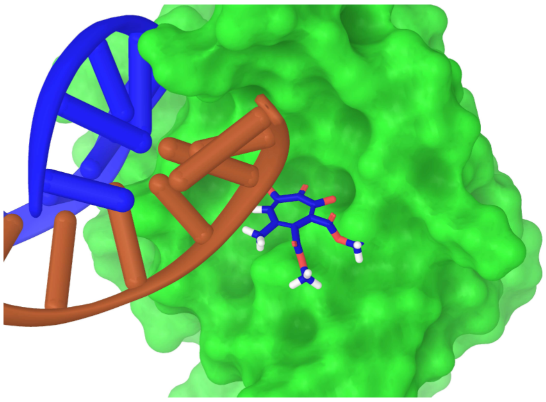


Figure 9. Homology model of the RNase H domain of HIV-1 RT (green surface) bound to the RNA/DNA hybrid substrate (blue and brown ladders) and ligand 1 (blue carbon atoms).

■ CONCLUSIONS

In this work, we have conducted a large-scale binding free energy-based virtual screening campaign of a library of 77 α-hydroxytropolone derivatives against the active site of the RNase H active site of the RT enzyme of HIV-1. Lead compounds emerging from the screening process were subjected to targeted optimization, which resulted in the identification of compounds with significantly increased potency. The work involved 280 individual absolute binding free energy calculations when multiple protonation states, multiple binding modes, and individual rotamer states of each ligand were considered. The computational effort for this work (over 1 million aggregate CPU hours) has been made feasible by automated protocols and the use of heterogeneous high performance computing resources at XSEDE and the WEB computational grid at Brooklyn College. The thermodynamic and structural measures obtained in this work rationalize a series of characteristics of this system useful for guiding future synthetic and biochemical efforts. Specifically, our calculations support the model that the nucleic acid substrate limits the range of binding modes available to the inhibitors and that it likely plays other key roles in the recruitment of inhibitors to the active site. The computational model further confirms that small or rigid side groups at R2, which point mostly to the solvent, have minor effects on affinity. On the other hand, larger and more flexible R2 side groups can lead to intramolecular hydrophobic interactions, which significantly decrease affinity. Substitutions at the R3 position emerge as more promising. Specific side groups here can have the effect of recruiting ARG 557 (which is otherwise disordered) to form a hydrophobically shielded ionic interaction with the ligand. Further targeted modifications at R3 (exemplified by ligand 81, Figure 8) are predicted to also induce the recruitment of HIS 539, a key residue involved in the nuclease reaction, forming a complex stabilized by a network of hydrogen bond, hydrophobic and $\pi-\pi$ interactions. Future synthetic and biochemical work will focus on exploiting the ligand design strategies emerging from this work. Overall, the results presented here show that modern binding free energy models, when intelligently integrated into interdisciplinary chemistry efforts, can offer useful insights and contribute in useful ways to early-stage drug-discovery programs.

■ ASSOCIATED CONTENT

Ⓢ Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acsomega.6b00123.

Chemical structures of the 279 ligand models investigated, ranked by ligand identification number (ligand-ID) (Table S1) (PDF)

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: egallicchio@brooklyn.cuny.edu.

Present Address

‖Department of Chemistry, CUNY Brooklyn College, 2900 Bedford Avenue, Brooklyn, New York 11210, United States (E.G.).

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

E.G. acknowledges support from the National Science Foundation (SI2-SSE 1440665). R.M. acknowledges support from the National Institutes of Health (SC1-GM111158). REMD simulations were carried out on the Supermic cluster of

XSEDE (supported by project TG-MCB150001 to E.G. and B.Z.) and the WEB computational grid at Brooklyn College of the City University of New York. The authors acknowledge invaluable computing support from James Roman and John Stephen at Brooklyn College.

■ REFERENCES

(1) Sarafianos, S. G.; Marchand, B.; Das, K.; Himmel, D. M.; Parniak, M. A.; Hughes, S. H.; Arnold, E. *J. Mol. Biol.* **2009**, *385*, 693−713.
(2) Tisdale, M.; Schulze, T.; Larder, B. A.; Moelling, K. *J. Gen. Virol.* **1991**, *72*, 59−66.
(3) Felts, A. K.; Labarge, K.; Bauman, J. D.; Patel, D. V.; Himmel, D. M.; Arnold, E.; Parniak, M. A.; Levy, R. M. *J. Chem. Inf. Model.* **2011**, *51*, 1986−1998.
(4) Himmel, D. M.; Maegley, K. A.; Pauly, T. A.; Bauman, J. D.; Das, K.; Dharia, C.; Clark, A. D., Jr.; Ryan, K.; Hickey, M. J.; Love, R. A.; Hughes, S. H.; Bergqvist, S.; Arnold, E. *Structure* **2009**, *17*, 1625−1635.
(5) Himmel, D.; Sarafianos, S.; Dharmasena, S.; Hossain, M.; McCoy-Simandle, K.; Ilina, T.; Clark, A., Jr.; Knight, J.; Julias, J.; Clark, P.; Krogh-Jespersen, K.; Levy, R.; Hughes, S.; Parniak, M.; Arnold, E. *ACS Chem. Biol.* **2006**, *1*, 702−712.
(6) Shaw-Reid, C.; Munshi, V.; Graham, P.; Wolfe, A.; Witmer, M.; Danzeisen, R.; Olsen, D.; Carroll, S.; Embrey, M.; Wai, J.; Miller, M.; Cole, J.; Hazuda, D. *J. Biol. Chem.* **2003**, *278*, 2777−2780.
(7) Su, H.-P.; Yan, Y.; Prasad, G. S.; Smith, R. F.; Daniels, C. L.; Abeywickrema, P. D.; Reid, J. C.; Loughran, H. M.; Kornienko, M.; Sharma, S.; Grobler, J. A.; Xu, B.; Sardana, V.; Allison, T. J.; Williams, P. D.; Darke, P. L.; Hazuda, D. J.; Munshi, S. *J. Virol.* **2010**, *84*, 7625−7633.
(8) Costi, R.; Métifiot, M.; Chung, S.; Crucitti, G. C.; Maddali, K.; Pescatori, L.; Messore, A.; Madia, V. N.; Pupo, G.; Scipione, L.; Tortorella, S.; Saverio Di Leva, F.; Cosconati, S.; Marinelli, L.; Novellino, E.; Le Grice, S. F. J.; Corona, A.; Pommier, Y.; Marchand, C.; Di Santo, R. *J. Med. Chem.* **2014**, *57*, 3223−3234.
(9) Himmel, D. M.; Myshakina, N. S.; Ilina, T.; Van Ry, A.; Ho, W. C.; Parniak, M. A.; Arnold, E. *J. Mol. Biol.* **2014**, *426*, 2617−2631.
(10) Esposito, F.; Tramontano, E. *Antiviral Chem. Chemother.* **2014**, *23*, 129−144.
(11) Lansdon, E.; Liu, Q.; Leavitt, S.; Balakrishnan, M.; Perry, J.; Lancaster-Moyer, C.; Kutty, N.; Liu, X.; Squires, N.; Watkins, W.; Kirschberg, T. *Antimicrob. Agents Chemother.* **2011**, *55*, 2905−2915.
(12) Fowler, B. J.; Gelfand, B. D.; Kim, Y.; Kerur, N.; Tarallo, V.; Hirano, Y.; Amarnath, S.; Fowler, D. H.; Radwan, M.; Young, M. T.; Pittman, K.; Kubes, P.; Agarwal, H. K.; Parang, K.; Hinton, D. R.; Bastos-Carvalho, A.; Li, S.; Yasuma, T.; Mizutani, T.; Yasuma, R.; Wright, C.; Ambati, J. *Science* **2014**, *346*, 1000−1003.
(13) Gallant, J.; Gerondelis, P.; Wainberg, M.; Shulman, N.; Haubrich, R.; St Clair, M.; Lanier, E.; Hellmann, N.; Richman, D. *Antiviral Ther.* **2003**, *8*, 489−506.
(14) Barreca, M. L.; Ferro, S.; Rao, A.; Luca, L. D.; Zappalá, M.; Monforte, A.-M.; Debyser, Z.; Witvrouw, M.; Chimirri, A. *J. Med. Chem.* **2005**, *48*, 7084−7088.
(15) McColl, D.J.; Chen, X. *Antiviral Res.* **2010**, *85*, 101−118.
(16) Davies, J.; Hostomska, Z.; Hostomsky, Z.; Jordan, S.; Matthews, D. *Science* **1991**, *252*, 88−95.
(17) Chung, S.; Himmel, D. M.; Jiang, J.-K.; Wojtak, K.; Bauman, J. D.; Rausch, J. W.; Wilson, J. A.; Beutler, J. A.; Thomas, C. J.; Arnold, E.; Le Grice, S. F. J. *J. Med. Chem.* **2011**, *54*, 4462−4473.
(18) Murelli, R. P.; D'Erasmo, M. P.; Hirsch, D. R.; Meck, C.; Masaoka, T.; Wilson, J. A.; Zhang, B.; Pal, R. K.; Gallicchio, E.; Beutler, J. A.; Grice, S. F. J. L. *Med. Chem. Commun.* **2016**, advance article.
(19) Goodsell, D. S.; Morris, G. M.; Olson, A. J. *J. Mol. Recognit.* **1996**, *9*, 1−5.
(20) Shoichet, B. K.; McGovern, S. L.; Wei, B.; Irwin, J. J. *Curr. Opin. Chem. Biol.* **2002**, *6*, 439−446.
(21) Forli, S.; Olson, A. J. *J. Med. Chem.* **2012**, *55*, 623−638.
(22) Perryman, A. L.; Santiago, D. N.; Forli, S.; Santos-Martins, D.; Olson, A. J. *J. Comput.-Aided Mol. Des.* **2014**, *28*, 429.
(23) Hogues, H.; Sulea, T.; Purisima, E. O. *J. Chem. Inf. Model.* **2016**, *56*, 955−964.
(24) Jorgensen, W. L. *Science* **2004**, *303*, 1813−1818.
(25) Wang, L.; Wu, Y.; Deng, Y.; Kim, B.; Pierce, L.; Krilov, G.; Lupyan, D.; Robinson, S.; Dahlgren, M. K.; Greenwood, J.; Romero, D. L.; Masse, C.; Knight, J. L.; Steinbrecher, T.; Beuming, T.; Damm, W.; Harder, E.; Sherman, W.; Brewer, M.; Wester, R.; Murcko, M.; Frye, L.; Farid, R.; Lin, T.; Mobley, D. L.; Jorgensen, W. L.; Berne, B. J.; Friesner, R. A.; Abel, R. *J. Am. Chem. Soc.* **2015**, *137*, 2695−2703.
(26) Boresch, S.; Tettinger, F.; Leitgeb, M.; Karplus, M. *J. Phys. Chem. B* **2003**, *107*, 9535−9551.
(27) Young, T.; Abel, R.; Kim, B.; Berne, B. J.; Friesner, R. A. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 808−813.
(28) Zhou, H.-X.; Gilson, M. K. *Chem. Rev.* **2009**, *109*, 4092−4107.
(29) Gallicchio, E.; Lapelosa, M.; Levy, R. M. *J. Chem. Theory Comput.* **2010**, *6*, 2961−2977.
(30) Gallicchio, E.; Levy, R. M. *Curr. Opin. Struct. Biol.* **2011**, *21*, 161−166.
(31) Gallicchio, E.; Levy, R. M. *Adv. Prot. Chem. Struct. Biol.* **2011**, *85*, 27−80.
(32) Gallicchio, E.; Chen, H.; Chen, H.; Fitzgerald, M.; Gao, Y.; He, P.; Kalyanikar, M.; Kao, C.; Lu, B.; Niu, Y.; Pethe, M.; Zhu, J.; Levy, R. M. *J. Comput.-Aided Mol. Des.* **2015**, *29*, 315−325.
(33) Gallicchio, E.; Deng, N.; He, P.; Perryman, A. L.; Santiago, D. N.; Forli, S.; Olson, A. J.; Levy, R. M. *J. Comput.-Aided Mol. Des.* **2014**, *28*, 475−490.
(34) Piettre, S. R.; Ganzhorn, A.; Hoflack, J.; Islam, K.; Hornsperger, J.-M. *J. Am. Chem. Soc.* **1997**, *119*, 3201−3204.
(35) Budihas, S.; Gorshkova, I.; Gaidamakov, S.; Wamiru, A.; Bona, M.; Parniak, M.; Crouch, R.; McMahon, J.; Beutler, J.; Le Grice, S. *Nucleic Acids Res.* **2005**, *33*, 1249−1256.
(36) Klumpp, K.; Hang, J. Q.; Rajendran, S.; Yang, Y.; Derosier, A.; Wong Kai In, P.; Overton, H.; Parkes, K. E. B.; Cammack, N.; Martin, J. A. *Nucleic Acids Res.* **2003**, *31*, 6852−6859.
(37) Gallicchio, E.; Paris, K.; Levy, R. M. *J. Chem. Theory Comput.* **2009**, *5*, 2544−2564.
(38) Gallicchio, E.; Levy, R. M.; Parashar, M. *J. Comput. Chem.* **2008**, *29*, 788−794.
(39) Hamelberg, D.; McCammon, J. A. *J. Am. Chem. Soc.* **2004**, *126*, 7683−7689.
(40) Madhavi Sastry, G.; Adzhigirey, M.; Day, T.; Annabhimoju, R.; Sherman, W. *J. Comput.-Aided Mol. Des.* **2013**, *27*, 221−234.
(41) Poongavanam, V.; Steinmann, C.; Kongsted, J. *PLoS One* **2014**, *9*, e98659.
(42) Zhao, H.; Lin, Z.; Lynn, A. Y.; Varnado, B.; Beutler, J. A.; Murelli, R. P.; Grice, S. F. J. L.; Tang, L. *Nucleic Acids Res.* **2015**, *43*, 11003−11016.
(43) Li, P.; Song, L. F.; Kenneth, M.; Merz, J. *J. Phys. Chem. B* **2015**, *119*, 883−895.
(44) Gardner, J.; Baron, G.; MacLean, H. *Can. J. Chem.* **1957**, *35*, 1039−1048.
(45) Martin, S. F.; Follows, B. C.; Hergenrother, P. J.; Franklin, C. L. *J. Org. Chem.* **2000**, *65*, 4509−4514.
(46) Greenwood, J. R.; Calkins, D.; Sullivan, A. P.; Shelley, J. C. *J. Comput.-Aided Mol. Des.* **2010**, *24*, 591−604.
(47) Mobley, D. L.; Chodera, J. D.; Dill, K. A. *J. Chem. Phys.* **2006**, *125*, No. 084902, DOI: 10.1063/1.2221683.
(48) Lapelosa, M.; Gallicchio, E.; Levy, R. M. *J. Chem. Theory Comput.* **2012**, *8*, 47−60.
(49) Gilson, M. K.; Given, J. A.; Bush, B. L.; McCammon, J. A. *Biophys. J.* **1997**, *72*, 1047−1069.
(50) Gallicchio, E.; Levy, R. M. *J. Comput.-Aided Mol. Des.* **2012**, *26*, 505−516.
(51) Gallicchio, E. *Comput. Mol. Biosci.* **2012**, *2*, 7−22.
(52) Tan, Z.; Gallicchio, E.; Lapelosa, M.; Levy, R. M. *J. Chem. Phys.* **2012**, *136*, No. 144102, DOI: 10.1063/1.3701175.

(53) Wickstrom, L.; He, P.; Gallicchio, E.; Levy, R. M. *J. Chem. Theory Comput.* **2013**, *9*, 3136−3150.
(54) Pohorille, A.; Jarzynski, C.; Chipot, C. *J. Phys. Chem. B* **2010**, *114*, 10235−10253.
(55) Shirts, M. R.; Chodera, J. D. *J. Chem. Phys.* **2008**, *129*, No. 124105, DOI: 10.1063/1.2978177.
(56) Feig, M.; Brooks, C. *Curr. Opin. Struct. Biol.* **2004**, *14*, 217−224.
(57) Qiu, D.; Shenkin, P. S.; Hollinger, F. P.; Still, C. W. *J. Phys. Chem. A* **1997**, *101*, 3005−3014.
(58) Gallicchio, E.; Levy, R. *J. Comput. Chem.* **2004**, *25*, 479−499.
(59) Mobley, D. L.; Dill, K. A. *Structure* **2009**, *17*, 489−498.
(60) Kaminski, G. A.; Friesner, R. A.; Tirado-Rives, J.; Jorgensen, W. L. *J. Phys. Chem. B* **2001**, *105*, 6474−6487.
(61) Banks, J.; Beard, J.; Cao, Y.; Cho, A.; Damm, W.; Farid, R.; Felts, A.; Halgren, T.; Mainz, D.; Maple, J.; Murphy, R.; Philipp, D.; Repasky, M.; Zhang, L.; Berne, B.; Friesner, R.; Gallicchio, E.; Levy, R. *J. Comput. Chem.* **2005**, *26*, 1752−1780.
(62) Gallicchio, E.; Xia, J.; Flynn, W. F.; Zhang, B.; Samlalsingh, S.; Mentes, A.; Levy, R. M. *Comput. Phys. Commun.* **2015**, *196*, 236−246.
(63) Sarafianos, S. G.; Das, K.; Tantillo, C.; Clark, A. D.; Ding, J.; Whitcomb, J. M.; Boyer, P. L.; Hughes, S. H.; Arnold, E. *EMBO J.* **2001**, *20*, 1449−1461.
(64) Bauman, J. D.; Patel, D.; Dharia, C.; Fromer, M. W.; Ahmed, S.; Frenkel, Y.; Vijayan, R. S. K.; Eck, J. T.; Ho, W. C.; Das, K.; Shatkin, A. J.; Arnold, E. *J. Med. Chem.* **2013**, *56*, 2738−2746.
(65) Lapkouski, M.; Tian, L.; Miller, J. T.; Le Grice, S. F.; Yang, W. *Nat. Struct. Mol. Biol.* **2013**, *20*, 230−236.
(66) Biebricher, A. S.; Heller, I.; Roijmans, R. F. H.; Hoekstra, T. P.; Peterman, E. J. G.; Wuite, G. J. L. *Nat. Commun.* **2015**, *6*, 7304.
(67) Christen, M. T.; Menon, L.; Myshakina, N. S.; Ahn, J.; Parniak, M. A.; Ishima, R. *Chem. Biol. Drug Des.* **2012**, *80*, 706−716.
(68) Gustafson, J. L.; Lim, D.; Miller, S. J. *Science* **2010**, *328*, 1251−1255.
(69) Friesner, R. A.; Banks, J. L.; Murphy, R. B.; Halgren, T. A.; Klicic, J. J.; Mainz, D. T.; Repasky, M. P.; Knoll, E. H.; Shelley, M.; Perry, J. K.; Shaw, D. E.; Francis, P.; Shenkin, P. S. *J. Med. Chem.* **2004**, *47*, 1739−1749.
(70) Cosconati, S.; Forli, S.; Perryman, A. L.; Harris, R.; Goodsell, D. S.; Olson, A. J. *Expert Opin. Drug Discovery* **2010**, *5*, 597−607.
(71) Singh, N.; Warshel, A. *Proteins* **2010**, *78*, 1705−1723.
(72) Poongavanam, V.; Olsen, J. M. H.; Kongsted, J. *Integr. Biol.* **2014**, *6*, 1010−1022.